

From the Peripheral Vascular Surgery Society

# Accuracy of administrative data versus clinical data to evaluate carotid endarterectomy and carotid stenting

Rodney P. Bensley, MD, Shunsuke Yoshida, MD, Ruby C. Lo, MD, Margriet Fokkema, MD, Allen D. Hamdan, MD, Mark C. Wyers, MD, Elliot L. Chaikof, MD, PhD, and Marc L. Schermerhorn, MD, *Boston, Mass*

**Objective:** Administrative data have been used to compare carotid endarterectomy (CEA) and carotid artery stenting (CAS). However, there are limitations in defining symptom status, Centers for Medicare and Medicaid Services high-risk status, as well as complications. Therefore, we did a direct comparison between administrative data and physician chart review as well as between data collected for the National Surgical Quality Improvement Program (NSQIP) and physician chart review for CEA and CAS.

**Methods:** We performed an outcomes analysis on all CEA and CAS procedures from 2005 to 2011. We obtained International Classification of Diseases, Ninth Revision diagnosis codes from hospital discharge records regarding symptom status, high-risk status, and perioperative stroke. We also obtained data on all CEA patients submitted to NSQIP over the same time period. One of the study authors (R.B.) then performed a chart review of the same patients to determine symptom status, high-risk status, and perioperative strokes and the results were compared.

**Results:** We identified 1342 patients who underwent CEA or CAS between 2005 and 2011 and 392 patients who underwent CEA that were submitted to NSQIP. Administrative data identified fewer symptomatic patients (17.0% vs 34.0%), physiologic high-risk patients (9.3% vs 23.0%), and anatomic high-risk patients (0% vs 15.2%). Although administrative data identified a similar proportion of perioperative strokes (1.9% vs 2.0%), this was due to the fact that these data identified eight false positive and nine false negative perioperative strokes. NSQIP data identified more symptomatic patients compared with chart review (44.1% vs 30.3%), fewer physiologic high-risk patients (13.0% vs 18.6%), fewer anatomic high-risk patients (0% vs 6.6%), and a similar proportion of perioperative strokes (1.5% vs 1.8%, only one false negative stroke and no false positives).

**Conclusions:** Administrative data are unreliable for determining symptom status, high-risk status, and perioperative stroke and should not be used to analyze CEA and CAS. NSQIP data do not adequately identify high-risk patients, but do accurately identify perioperative strokes and to a lesser degree, symptom status. (*J Vasc Surg* 2013;58:412-20.)

Administrative data have been utilized to compare carotid endarterectomy (CEA) and carotid artery stenting (CAS).<sup>1-3</sup> Large databases populated with administrative data are valuable research tools. However, the accuracy of administrative data in determining pre-existing disease, symptom status, high-risk status, and perioperative complications has been questioned.<sup>4</sup> Studies utilizing the

Nationwide Inpatient Sample report that 90%-97% of the patients undergoing carotid revascularization are asymptomatic,<sup>1-3</sup> whereas multicenter studies and retrospective reviews report that only 56%-72% of patients are asymptomatic.<sup>5-8</sup> International Classification of Diseases, Ninth Revision (ICD-9) codes are nonspecific and imprecise. They do not specify the extent of disease within a diagnosis (eg, class III-IV congestive heart failure [CHF]), do not provide laterality of disease, and lack the temporal timing of onset. These limitations with ICD-9 codes limit the ability to distinguish pre-existing disease from new disease and perioperative complications.

Other large datasets that utilize clinical data, such as the National Surgical Quality Improvement Program (NSQIP), may provide a more accurate measurement of symptom status, high-risk status, and perioperative complications. NSQIP is not reliant on hospital coders. Trained clinical nurse reviewers input data prospectively. However, the current iteration of NSQIP does not have procedure-specific comorbidities or complications. For example, a history of transient ischemic attack (TIA) or stroke are obtained but may have occurred more than 6 months before surgery or may have been contralateral to the carotid undergoing treatment. Similarly, there is

From the Department of Surgery, Division of Vascular and Endovascular Surgery, Beth Israel Deaconess Medical Center and Harvard Medical School. This work was supported by the NIH T32 Harvard-Longwood Research Training in Vascular Surgery grant HL007734.

Author conflict of interest: Dr Schermerhorn is a consultant for Endologix, Medtronic, and Boston Scientific.

Presented at the plenary session of the Thirty-seventh Annual Meeting of the Peripheral Vascular Surgery Society, National Harbor, Md, June 6, 2012.

Reprint requests: Marc L. Schermerhorn, MD, Department of Surgery, Division of Vascular and Endovascular Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, 110 Francis St, Ste 5B, Boston, MA 02215 (e-mail: [mscherm@bidmc.harvard.edu](mailto:mscherm@bidmc.harvard.edu)).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214/\$36.00

Copyright © 2013 by the Society for Vascular Surgery.

<http://dx.doi.org/10.1016/j.jvs.2013.01.010>

no measure of disease severity. The ability to accurately determine symptom status and high-risk status with NSQIP has not been studied. The purpose of this study is to determine the accuracy with which administrative data and NSQIP data capture symptom status, high-risk status, and perioperative complications compared with chart review by a trained physician.

## METHODS

**Patients.** We obtained hospital administrative discharge data on all patients undergoing CEA or CAS from January 1, 2005 to December 31, 2011 at our institution. Patients were identified using ICD-9 procedure codes for CEA (38.12) or CAS (00.63). We also identified all patients undergoing CEA at our institution that were submitted to NSQIP. We used our institution's administrative discharge data and prospectively collected clinical data submitted to NSQIP, to determine symptom status, high-risk status, and perioperative stroke and compared these results with physician chart review. This study was approved by the Institutional Review Board at Beth Israel Deaconess Medical Center and Harvard Medical School.

**Hospital administrative discharge data.** We identified patients undergoing CEA or CAS. Patients having coronary artery bypass (CABG) (36.11-36.16) or cardiac valve repair (35.X) during the same hospitalization were identified. All ICD-9 diagnosis codes associated with each patient's discharge record were reviewed to determine symptom status and physiologic and anatomic high-risk status. Physiologic high-risk variables include age >80 years, CHF class III/IV, left ventricular ejection fraction (LVEF) <30%, unstable angina, myocardial infarction (MI) within 30 days, severe lung disease (forced expiratory volume in 1 second <30% of predicted or home oxygen), hemodialysis, and CABG or valve repair within 30 days. Anatomic high-risk variables include contralateral laryngeal nerve palsy, restenosis after prior CEA, radiation therapy to the neck, a high lesion, prior neck surgery, and contralateral internal carotid artery occlusion. Patients were considered symptomatic if they had ICD-9 codes for TIA (435.X, 781.4, V12.54), amaurosis fugax/retinal vascular occlusion (362.3X, 368.12), or stroke (433.11, 433.31, 433.91, 434.01, 434.11, 434.91). Patients without any of these diagnosis codes were considered asymptomatic. As has been previously reported, ICD-9 diagnosis codes were used to identify comorbid conditions that would qualify the patient for Centers for Medicare and Medicaid Services (CMS) physiologic high-risk status<sup>3</sup> (Table I). Patients without these conditions were considered non-high risk. Perioperative strokes were identified using ICD-9 code 997.02 (iatrogenic cerebrovascular infarction or hemorrhage).

**Prospectively collected NSQIP data.** NSQIP is a national multi-institutional, risk-adjusted, prospectively collected clinical database created to facilitate quality control review of outcomes. The registry collects information on demographics, comorbidities, intraoperative information, and 30-day perioperative outcomes. Specially

trained clinical nurse reviewers, trained by the ACS, input data at each participating institution using standardized definitions. We obtained a list of all patients undergoing CEA at our institution whose data were submitted to NSQIP. We calculated the proportion of patients that were symptomatic, physiologic and anatomic high-risk, and those with perioperative strokes. Physiologic high-risk variables utilized in NSQIP include age >80 years, active CHF, MI within 6 months, angina within 1 month, and hemodialysis.<sup>9</sup> Patients were considered symptomatic if they had a history of TIA or stroke.

**Physician chart review.** One of the study authors (R.B.) performed a thorough chart review of the same patients identified using hospital administrative data and the patients submitted to NSQIP. The senior author clarified all questions encountered during chart review. Symptom status, physiologic and anatomic high-risk status, and perioperative strokes were identified. Symptomatic was defined as ipsilateral carotid territory TIA or stroke within 6 months of their carotid procedure. Discrepancies in the identification of symptom status, high-risk status, and perioperative strokes were identified and recorded. The majority of patients were not examined by a neurologist unless there was a suspicion for a perioperative stroke or if they were enrolled in a clinical trial requiring neurologic evaluation.

**Defining high-risk variables.** CMS has specific criteria defining high-risk for CEA.<sup>10</sup> When performing chart review, all CMS guidelines were adhered to when identifying both physiologic and anatomic high-risk patients. However, ICD-9 coding lacks the level of detail and specificity outlined by CMS. Furthermore, there are no ICD-9 codes for anatomic high-risk variables so these were unable to be assessed with administrative data. NSQIP has specific definitions of comorbidities that differ from CMS<sup>11</sup> and lacks anatomic details. Table I lists high-risk variables (physiologic and anatomic) and how they are identified and defined by CMS, administrative data, and NSQIP, respectively.

**Statistical analysis.** We calculated the proportion of patients who were considered symptomatic, physiologic high-risk, anatomic high-risk, and who experienced a perioperative stroke by administrative data and NSQIP data, respectively, and then compared them with physician chart review of the same patients. We then calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and kappa coefficient of the ability of administrative data and NSQIP data to properly identify symptom status, high-risk status, and perioperative strokes as compared with chart review (Fig 1 shows the 2 × 2 table used to calculate the sensitivity, specificity, PPV, and NPV). Sensitivity measures the accuracy of identifying a variable (ie, symptomatic, high-risk) as being present in patients in whom it truly is present. Specificity measures the accuracy of identifying a variable as being absent in patients who do not have the variable. PPV measures the proportion of patients identified as having a variable that truly have the variable. NPV measures the proportion of patients

**Table I.** Physiologic and anatomic high-risk variables and how they are defined by the CMS, administrative data, and the NSQIP database

<i>CMS definitions</i>	<i>Administrative data</i>	<i>NSQIP definitions</i>
Physiologic high risk		
Age >80 years	Age >80 years	Age >80 years
Recent MI within 30 days	412 <sup>b</sup>	History of a non-Q wave or a Q wave MI within 6 months of surgery
Unstable angina	411.1 <sup>b</sup>	Angina within 30 days of surgery
CHF class III or IV	428 <sup>b</sup>	Newly diagnosed CHF or chronic CHF with new symptoms within 30 days of surgery
LVEF <30%	Unavailable	Unavailable
Hemodialysis	585.3-585.9, 586, V42.0, V45.1-V45.12, V56.0 - V56.32, V56.8 <sup>b</sup>	Acute or chronic renal failure requiring peritoneal dialysis, hemodialysis, hemofiltration, hemodiafiltration, or ultrafiltration within 2 weeks of surgery
Severe pulmonary disease <sup>a</sup>	490-492.8, 493, 494-494.1, 495-505, 506.4 <sup>b</sup>	Severe COPD <sup>c</sup>
CABG/valve repair within 30 days	36.11-36.16, 35 <sup>b</sup>	CABG/valve repair performed at the same time as CEA
Anatomic high risk		
Contralateral ICA occlusion	Unavailable	Unavailable
Restenosis after CEA	Unavailable	Unavailable
Prior neck surgery	Unavailable	Unavailable
Contralateral laryngeal nerve palsy	Unavailable	Unavailable
History of neck irradiation	Unavailable	Unavailable
High or low lesion	Unavailable	Unavailable

CABG, Coronary artery bypass grafting; CEA, carotid endarterectomy; CHF, congestive heart failure; CMS, Centers for Medicare and Medicaid Services; COPD, chronic obstructive pulmonary disease; ICA, internal carotid artery; ICD-CM, International Classification of Diseases, Ninth Revision-clinically modified; LVEF, left ventricular ejection fraction; MI, myocardial infarction; NSQIP, National Surgical Quality Improvement Program.

<sup>a</sup>Forced expiratory volume in 1 second <30% of predicted or home oxygen.

<sup>b</sup>ICD-9-CM codes.

<sup>c</sup>Functional disability from COPD, hospitalization for treatment of COPD, chronic bronchodilator therapy, or a forced expiratory volume in 1 second <75% of predicted.

		Variable as identified on chart review	
		Present	Not Present
Variable as identified by administrative or NSQIP data	Present	True Positive	False Positive
	Not Present	False Negative	True Negative

**Fig 1.** The 2 × 2 table used to calculate sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). NSQIP, National Surgical Quality Improvement Program.

identified as not having a variable that truly do not have that variable. Kappa is a measure of agreement beyond chance between two databases or methods of identifying data. Kappa values >0.75 indicate excellent agreement,

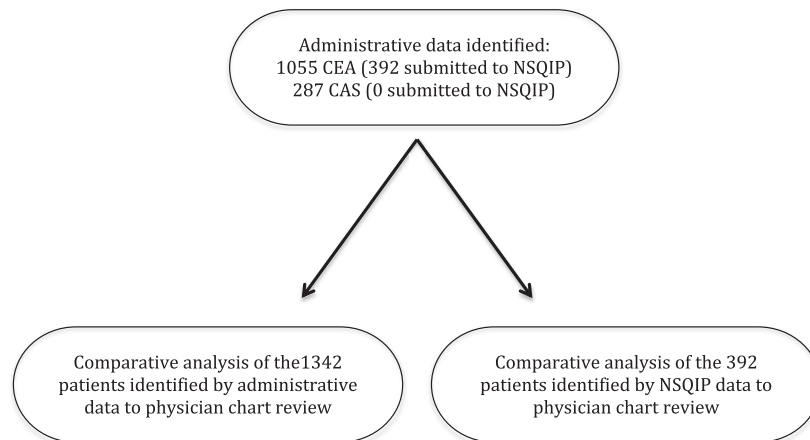
0.40-0.75 fair to good agreement, and values <0.4 poor agreement. All statistical tests were performed using STATA 12 software (StataCorp, College Station, Tex).

## RESULTS

Administrative data identified 1342 patients who underwent carotid revascularization (1055 CEA and 287 CAS) at our institution from 2005 to 2011. We identified 392 CEA patients submitted to NSQIP, all of whom were included in the 1055 CEA patients identified by administrative data. A chart review was then performed of the 1342 patients identified with administrative data and the 392 patients submitted to NSQIP to make direct comparisons to each method (Fig 2).

### Comparison of administrative data and physician chart review

Comparing administrative data to chart review, administrative data identified fewer symptomatic patients (17.0% vs 34.0%), fewer physiologic high-risk patients (9.3% vs 23.0%), fewer anatomic high-risk patients (0% vs 15.2%), and a similar proportion of perioperative strokes (1.9% vs 2.0%). When examining individual high-risk variables, administrative data overestimated recent MI, CHF, hemodialysis, and severe pulmonary disease, whereas underestimating unstable angina, LVEF <30%, and CABG/valve



**Fig 2.** Flow diagram depicting all patients captured by administrative data, National Surgical Quality Improvement Program (NSQIP) data, and physician chart review. CAS, Carotid artery stenting; CEA, carotid endarterectomy.

**Table II.** Comparison of outcomes analysis using administrative data and physician chart review in 1342 patients who underwent CEA or CAS between 2005 and 2011

	Administrative data		Physician chart review	
	No.	%	No.	%
Symptomatic	228	17.0	456	34.0
Physiologic high risk	125	9.3	308	23.0
Age >80 years	265	19.7	265	19.7
Recent MI	174	13.0	9	0.7
Unstable angina	1	0.0	16	1.2
CHF class III or IV	125	9.3	10	0.7
LVEF <30%	-	-	35	2.6
Hemodialysis	108	8.0	13	1.0
Severe pulmonary disease <sup>a</sup>	250	18.6	15	1.1
CABG/valve repair within 30 days	32	2.4	48	3.6
Anatomic high risk	-	-	204	15.2
Contralateral ICA occlusion	-	-	82	6.1
Restenosis after CEA	-	-	108	8.0
Prior neck surgery	-	-	41	3.1
Contralateral laryngeal nerve palsy	-	-	2	0.1
History of neck irradiation	-	-	29	2.2
High or low lesion	-	-	29	2.2
Perioperative stroke	26	1.9	27	2.0

CABG, Coronary artery bypass grafting; CEA, carotid endarterectomy; CHF, congestive heart failure; ICA, internal carotid artery; LVEF, left ventricular ejection fraction; MI, myocardial infarction.

<sup>a</sup>Forced expiratory volume in 1 second <30% of predicted or home oxygen.

repair (Table II). Administrative data are unable to identify anatomic high-risk variables using ICD-9 codes.

**Symptom status.** Administrative data identified 228 symptomatic patients, whereas chart review identified 456. However, there was poor overlap. Of 228 symptomatic patients identified by administrative data, only 167 were truly symptomatic. This means that 289 of the 456 symptomatic

patients identified by chart review were misclassified as being asymptomatic by administrative data. The sensitivity of administrative data for determining symptom status is 36.6%, specificity 93.1%, PPV 73.2%, and NPV 74.1%. The kappa is 0.34, indicating poor agreement between administrative data and chart review (Table III).

**High-risk status.** Administrative data identified 125 physiologic high-risk patients, whereas chart review identified 308. Only 60 of the 125 were truly physiologic high-risk, leaving 248 of the 308 misclassified by administrative data as being non-high risk. The sensitivity of administrative data for determining high-risk status is 19.5%, specificity 93.7%, PPV 48.0%, and NPV 79.6%. The kappa is 0.17, indicating poor agreement between administrative data and chart review. There was also poor concordance with the individual high-risk variables with the exception of age >80 years and CABG/valve repair (Table III). The kappa for age >80 and CABG/valve repair is 1.00 and 0.78, respectively, indicating excellent agreement. The kappa for the other high-risk variables ranges from -0.03 to 0.11 indicating poor agreement.

**Perioperative stroke.** Administrative data identified 26 perioperative strokes, whereas chart review identified 27. Of the 26 strokes identified by administrative data, 18 were true perioperative strokes and eight were false positive identifications. Nine perioperative strokes were missed entirely by administrative data. Table IV lists the explanations for the eight false positive and nine false negative perioperative strokes identified by administrative data. The sensitivity of administrative data for determining perioperative strokes is 66.7%, specificity 99.4%, PPV 69.2%, and NPV 99.3%. The kappa is 0.67 indicating fair to good agreement with chart review (Table III).

#### Comparison of NSQIP data and physician chart review

When comparing NSQIP to chart review, NSQIP identified more symptomatic patients (44.1% vs 30.3%), fewer physiologic high-risk patients (13.0% vs 18.6%),

**Table III.** Sensitivity, specificity, PPV, NPV, and kappa of administrative data compared with chart review

	<i>Sensitivity, %</i>	<i>Specificity, %</i>	<i>PPV, %</i>	<i>NPV, %</i>	<i>Kappa</i>
Symptomatic	36.6	93.1	73.2	74.1	0.34
Physiologic high risk	19.5	93.7	48.0	79.6	0.17
Age >80 years	100	100	100	100	1.00
MI within 30 days	33.3	55.4	1.7	97.2	-0.01
Unstable angina	6.3	100	100	96.2	0.11
CHF class III or IV	10.0	0.4	0.8	96.6	-0.03
LVEF <30%	-	-	-	-	-
Hemodialysis	61.5	73.6	7.4	98.2	0.08
Severe pulmonary disease <sup>a</sup>	46.7	35.5	2.8	94.4	-0.02
CABG/valve repair within 30 days	66.7	100	100	95.6	0.78
Perioperative stroke	66.7	99.4	69.2	99.3	0.67

CABG, Coronary artery bypass grafting; CHF, congestive heart failure; LVEF, left ventricular ejection fraction; MI, myocardial infarction; NPV, negative predictive value; PPV, positive predictive value.

<sup>a</sup>Forced expiratory volume in 1 second <30% of predicted or on home oxygen.

**Table IV.** Explanations for the eight false positive strokes and nine false negative strokes identified by administrative data

#### *Eight false positive perioperative strokes*

Four patients were falsely coded as having a perioperative stroke  
 One preoperative stroke was miscoded as being a postoperative stroke  
 Three strokes occurred after admission to the hospital for other reasons  
 One occurred after a laparoscopic cholecystectomy  
 One occurred after admission to the hospital for treatment of pneumonia  
 One occurred after admission to the hospital for lower extremity ischemia

#### *Nine false negative perioperative strokes*

Three were postdischarge strokes  
 Two were misclassified as preoperative strokes  
 Four strokes were not coded at all in the discharge data

fewer anatomic high-risk patients (0% vs 6.6%), and a similar proportion of perioperative strokes (1.5% vs 1.8%). When examining individual high-risk variables, NSQIP overestimated recent MI, unstable angina, CHF, and severe pulmonary disease, whereas underestimating LVEF <30% and hemodialysis (Table V). There was complete agreement in identifying patients >80 years and those who underwent CABG/valve repair. NSQIP is unable to identify anatomic high-risk variables.

**Symptom status.** NSQIP identified 173 symptomatic patients, whereas chart review identified 119. Of the 173 patients identified by NSQIP, 109 were truly symptomatic. This means that 10 of the 119 symptomatic patients identified by chart review were misclassified as being asymptomatic using NSQIP. The sensitivity of NSQIP for determining symptomatic status is 91.6%, specificity 76.6%, PPV 63.0%, and NPV 95.4%. The kappa of NSQIP is 0.60 indicating fair to good agreement between NSQIP and chart review (Table VI).

**High-risk status.** NSQIP identified 51 physiologic high-risk patients, whereas chart review identified 73. Only 10 of the 51 were truly physiologic high-risk patients leaving 63 of the 73 misclassified by NSQIP as being non-high risk. The sensitivity of NSQIP for determining high-risk status is 13.7%, specificity 98.4%, PPV 58.3%, and NPV 82.6%. The kappa of NSQIP is 0.01, indicating poor agreement between NSQIP and chart review. There was also poor concordance with the individual high-risk variables with the exception of age >80 years, hemodialysis, and CABG/valve repair (kappa of 1.00, 0.75, and 1.00, respectively, indicating excellent agreement) (Table VI). The kappa for the other high-risk variables ranges from -0.01 to 0.21 indicating poor agreement.

**Perioperative stroke.** NSQIP identified six perioperative strokes, whereas chart review identified seven. There was one false negative stroke and zero false positives with NSQIP. This one missed perioperative stroke was simply not captured by the person responsible for NSQIP data input. The sensitivity of NSQIP for determining perioperative stroke is 85.7%, specificity 100%, PPV 100%, and NPV 99.7%. The kappa of NSQIP is 0.92, indicating excellent agreement with chart review (Table VI).

## DISCUSSION

Many studies have used administrative data to study patients undergoing CEA or CAS.<sup>1-3</sup> The benefit of these databases is their large sample size, whereas the downside is their reliance on ICD-9 hospital discharge coding. With this study, we sought to examine the accuracy of ICD-9 coding compared with chart review, and we have found poor concordance between administrative and clinical data. Administrative data are poor at determining symptom status and high-risk status. Although perioperative stroke rates were similar between administrative data and chart review, administrative data identified the wrong patients and had poor accuracy. NSQIP data were also poor at determining high-risk status, but were accurate for determination of perioperative strokes and to a lesser extent symptom status.



**Table V.** Comparison of outcomes analysis using NSQIP data and physician chart review in 392 patients who underwent CEA between 2005 and 2011

	NSQIP data		Physician chart review	
	No.	%	No.	%
Symptomatic	173	44.1	119	30.3
Physiologic high risk	51	13.0	73	18.6
Age >80 years	76	19.4	76	19.4
Recent MI	7	1.8	2	0.5
Unstable angina	8	2.0	1	0.3
CHF class III or IV	3	0.8	0	0.0
LVEF <30%	-	-	4	1.0
Hemodialysis	3	0.8	5	1.3
Severe pulmonary disease <sup>a</sup>	41	10.4	4	1.0
CABG/valve repair within 30 days	3	0.8	3	0.8
Anatomic high risk	-	-	26	6.6
Contralateral ICA occlusion	-	-	20	5.1
Restenosis after CEA	-	-	15	3.8
Prior neck surgery	-	-	4	1.0
Contralateral laryngeal nerve palsy	-	-	0	0.0
History of neck irradiation	-	-	2	0.5
High or low lesion	-	-	2	0.5
Perioperative stroke	6	1.5	7	1.8

CABG, Coronary artery bypass grafting; CEA, carotid endarterectomy; CHF, congestive heart failure; ICA, internal carotid artery; LVEF, left ventricular ejection fraction; MI, myocardial infarction; NSQIP, National Surgical Quality Improvement Program.

<sup>a</sup>Forced expiratory volume in 1 second <30% of predicted or home oxygen.

Our study is the first to examine the accuracy of identifying symptom status and physiologic high-risk status with administrative data. Recently, the validity of prior studies utilizing administrative data to compare outcomes after CEA and CAS has been questioned.<sup>4</sup> Our results validate this question as we have found that administrative data cannot accurately determine symptom status. ICD-9 codes lack the specificity needed to accurately determine the extent and severity of a patient's comorbid condition, lack laterality of disease, and lack the temporal timing of onset. Because of this lack of detail, the use of ICD-9 codes to determine high-risk status may overestimate certain individual high-risk variables such as class III/IV CHF, recent MI, severe pulmonary disease, and hemodialysis. For example, when examining the rate of recent MI, administrative data identified 174 patients, whereas chart review identified nine. This discrepancy is due to the lack of specificity and timing of ICD-9 codes. ICD-9 diagnosis code 412 is used for myocardial infarction. However, it does not specify timing of the MI (within the past 30 days or remote). Chart review was able to determine that only nine patients had a MI within 30 days of their carotid procedure. Furthermore, there are no ICD-9 codes available to estimate anatomic high-risk status.

Other studies have examined the accuracy with which acute ischemic strokes are identified with administrative data, but not in the setting of CEA. Goldstein<sup>12</sup> reported that only 79% of patients with a primary discharge

diagnosis indicating cerebral infarction actually had an acute stroke and that 7% of patients with an ICD-9 code indicating the absence of cerebral infarction actually had an acute stroke. Reker et al<sup>13</sup> also showed that there is incredible variability in which ICD-9 code is reported to indicate an acute stroke and to the accuracy with which these codes actually represent new strokes. The rates of perioperative stroke were similar in our study, but this is misleading. Administrative data falsely identified patients as having a perioperative stroke and failed to identify others that did. In our study, three of 27 patients experienced their perioperative stroke after discharge from the hospital. Postdischarge events are not captured with hospital discharge data, and this highlights a very important limitation of most administrative datasets. Studies have shown that in-hospital data underestimates the true risk of perioperative stroke.<sup>5,6,14</sup> NSQIP, however, provides data on outcomes out to 30 days. We believe that true perioperative outcomes include 30-day outcomes and our study highlights that administrative data underestimates the true 30-day perioperative risk. In the setting of CEA, the analysis is further complicated by the distinction of preoperative stroke (indicating CEA for symptomatic disease) vs a postoperative complication, as stroke is a difficult endpoint to measure accurately with administrative data. Mortality, however, is a hard endpoint and can be accurately determined with administrative data.<sup>15,16</sup>

We found NSQIP to be poor at determining high-risk status most likely due to the different definitions NSQIP uses to define patient comorbid conditions. CMS has specific high-risk criteria and the NSQIP definition of these differs substantially. CMS is more specific and specifies a greater severity of disease and/or a shorter timeframe for an event to occur. Prior studies have used NSQIP to analyze CEA. One study by Kang et al<sup>9</sup> did not attempt to determine symptom status because NSQIP does not provide laterality (of symptoms or procedure) and only states if the patient has a history of TIA or stroke and does not provide timing of the event. Gupta et al<sup>17</sup> did stratify patients by symptom status. Similar to our analysis, they specified that a patient was symptomatic if they had a history of a prior TIA or stroke and found a similar proportion of symptomatic patients (43.5% vs 44.1%). This methodology overestimates the number of symptomatic patients as the sensitivity was 91.6% but the PPV was only 63.0%. This indicates that NSQIP is more accurate at identifying those patients who are truly symptomatic but also identifies patients in whom stroke or TIA was likely more than 6 months before their CEA procedure or contralateral, thus, making them asymptomatic. One perioperative stroke was missed by NSQIP, and no false positive strokes were identified. NSQIP is more accurate at determining perioperative strokes but remains poor at identifying high-risk patients and, therefore, has limited value for comparative effectiveness of CEA and CAS. These limitations of NSQIP have prompted a new iteration of NSQIP to be created, which is currently in use and now captures procedure-specific comorbidities and complications, which

**Table VI.** Sensitivity, specificity, PPV, NPV, and kappa of NSQIP data compared with chart review

	<i>Sensitivity, %</i>	<i>Specificity, %</i>	<i>PPV, %</i>	<i>NPV, %</i>	<i>Kappa</i>
Symptomatic	91.6	76.6	63.0	95.4	0.60
Physiologic high risk	13.7	87.1	19.6	81.5	0.01
Age >80 years	100	100	100	100	1.00
MI within 30 days	50.0	98.5	14.3	99.7	0.21
Unstable angina	0.0	98.0	0.0	99.7	-0.01
CHF class III or IV	0.0	99.2	0.0	100	-0.01
LVEF <30%	-	-	-	-	-
Hemodialysis	60.0	100	100	99.5	0.75
Severe pulmonary disease <sup>a</sup>	75.0	90.2	7.3	99.7	0.11
CABG/valve repair within 30 days	100	100	100	100	1.00
Perioperative stroke	85.7	100	100	99.7	0.92

CABG, Coronary artery bypass grafting; CHF, congestive heart failure; LVEF, left ventricular ejection fraction; MI, myocardial infarction; NPV, negative predictive value; NSQIP, National Surgical Quality Improvement Program; PPV, positive predictive value.

<sup>a</sup>Forced expiratory volume in 1 second <30% of predicted or on home oxygen.

should make NSQIP a better research and quality improvement tool in the future.<sup>18</sup>

Other large clinical databases such as the Vascular Study Group of New England (VSGNE) and the Vascular Quality Initiative (VQI) may be better designed to accurately study CEA and CAS. The VSGNE is a regional quality improvement registry that prospectively collects over 140 data points<sup>19</sup> and specifically identifies preoperative symptom status, laterality of disease and procedure, anatomic details, high-risk status, as well as the laterality of perioperative strokes. This level of clinical detail makes the VSGNE and VQI valuable research tools.

Medicare data may provide a higher level of accuracy compared with purely hospital discharge data since the timing of prior procedures (eg, CABG) could potentially be captured using longitudinal data. Additionally, Medicare may be useful in determining the date of onset of new neurologic symptoms or complications, thus, potentially improving the accuracy of determining symptom status and postprocedure strokes. However, knowledge of this timing would not be expected to improve the inherent limitations in the ability to identify postprocedure strokes or the severity of comorbid conditions that we identified in this analysis.

Because of the limitations associated with using administrative data, a present on admission (POA) indicator has been developed to improve coding accuracy. The goal of the POA indicator is to help determine if a diagnosis is a pre-existing disease, a new diagnosis, or a complication that occurred during the hospitalization. Prior studies have shown that the addition of a POA indicator more accurately characterized preexisting conditions from complications.<sup>20,21</sup> This should help in determining if a stroke is a perioperative complication or the reason for the procedure. However, the POA indicator will not aid in determining high-risk status, as the current ICD-9 codes are less specific than needed. Improved education of hospital coders may improve the accuracy and reliability of administrative data. ICD-10 codes should improve

diagnostic accuracy with regards to symptom status, high-risk status, and perioperative strokes, but the extent of this improvement cannot be determined until ICD-10 codes are fully implemented.

As stated earlier, administrative data have well known limitations. NSQIP is a clinical database; however, it is expensive and voluntary. NSQIP only captures a random selection of all CEAs performed at an institution and does not collect data on CAS. The preoperative characteristics and postoperative complications collected are not procedure specific, NSQIP has definitions of comorbid conditions that differ from CMS, and it does not capture anatomic high-risk criteria. NSQIP does specify a history of TIA or stroke but provides no details on the timing or laterality.

Physician chart review is susceptible to limitations as well. Despite thorough review, information can be overlooked and missed entirely. The accuracy of chart review is reliant upon complete and accurate documentation within the chart. Minor periprocedural strokes may be missed by the clinician or not entered into the medical record. As this is a retrospective review, not all patients were evaluated by a neurologist or had neuroimaging. The true "gold standard" for accurately determining comorbidities and complications is uncertain. However, we believe that physician chart review is the most accurate method available for retrospective analysis. Clearly, further comparisons will be needed to monitor the progress of each of these systems and the relative value of each for answering various clinical questions.

## CONCLUSIONS

Administrative data are poor at determining symptom status, high-risk status, and accurately detecting perioperative strokes after CEA and CAS. This is in large part due to the lack of specificity with ICD-9 diagnosis codes as they fail to provide information on the severity, laterality, and temporal onset of disease. However, although NSQIP does not adequately identify high-risk patients, it is more accurate at determining perioperative strokes and to a lesser

degree, symptom status. The new iteration of NSQIP, as it captures procedure-specific comorbid conditions and complications, will make this an even more robust database in the future.

#### AUTHOR CONTRIBUTIONS

Conception and design: RB, SY, RL, MF, AH, MW, EC, MS  
Analysis and interpretation: RB, MS

Data collection: RB, SY

Writing the article: RB, MS

Critical revision of the article: RB, SY, RL, MF, AH, MW, EC, MS

Final approval of the article: RB, SY, RL, MF, AH, MW, EC, MS

Statistical analysis: RB, SY

Obtained funding: Not applicable

Overall responsibility: MS

#### REFERENCES

1. Eslami MH, McPhee JT, Simons JP, Schanzer A, Messina LM. National trends in utilization and postprocedure outcomes for carotid artery revascularization 2005 to 2007. *J Vasc Surg* 2011;53:307-15.
2. Vogel TR, Dombrovsky VY, Haser PB, Scheirer JC, Graham AM. Outcomes of carotid artery stenting and endarterectomy in the United States. *J Vasc Surg* 2009;49:325-30.
3. Giles KA, Hamdan AD, Pomposelli FB, Wyers MC, Schermerhorn ML. Stroke and death after carotid endarterectomy and carotid artery stenting with and without high risk criteria. *J Vasc Surg* 2010;52:1497-504.
4. Hertzner NR. The Nationwide Inpatient Sample may contain inaccurate data for carotid endarterectomy and carotid stenting. *J Vasc Surg* 2012;55:263-6.
5. Sidawy AN, Zwolak RM, White RA, Siami FS, Schermerhorn ML, Sicard GA, et al. Risk-adjusted 30-day outcomes of carotid stenting and endarterectomy: results from the SVS Vascular Registry. *J Vasc Surg* 2009;49:71-9.
6. Goodney PP, Likosky DS, Cronenwett JL; Vascular Study Group of Northern New England. Factors associated with stroke or death after carotid endarterectomy in Northern New England. *J Vasc Surg* 2008;48:1139-45.
7. Hobson RW 2nd, Howard VJ, Roubin GS, Brott TG, Ferguson RD, Pompa JJ, et al. Carotid artery stenting is associated with increased complications in octogenarians: 30-day stroke and death rates in the CREST lead-in phase. *J Vasc Surg* 2004;40:1106-11.
8. Halm EA, Tuhim S, Wang JJ, Rockman C, Riles TS, Chassin MR. Risk factors for perioperative death and stroke after carotid endarterectomy: results of the New York Carotid Artery Surgery Study. *Stroke* 2009;40:221-9.
9. Kang JL, Chung TK, Lancaster RT, LaMuraglia GM, Conrad MF, Cambria RP. Outcomes after carotid endarterectomy: is there a high-risk population? A National Surgical Quality Improvement Program report. *J Vasc Surg* 2009;49:331-9.
10. Pub. 100-103 Medicare National Coverage Determinations. Department of Health and Human Services, Centers for Medicare and Medicaid Services; 2006.
11. American College of Surgeons, National Quality Improvement Program. Available at: [http://site.acsnsqip.org/wp-content/uploads/2012/03/ACS\\_NSQIP\\_Participant\\_User\\_Data\\_File\\_User\\_Guide.pdf](http://site.acsnsqip.org/wp-content/uploads/2012/03/ACS_NSQIP_Participant_User_Data_File_User_Guide.pdf). Accessed May 25, 2012.
12. Goldstein LB. Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke. *Stroke* 1998;29:1602-4.
13. Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A. Stroke: who's counting what? *J Rehabil Res Dev* 2001;38:281-9.
14. Fokkema M, Bensley RP, Lo RC, Hamdan AD, Wyers MC, Moll FL, et al. In-hospital versus postdischarge adverse events following carotid endarterectomy. *J Vasc Surg* 2013;57:1568-75.
15. Holt PJ, Poloniecki JD, Thompson MM. Multicentre study of the quality of a large administrative data set and implications for comparing death rates. *Br J Surg* 2012;99:58-65.
16. Wennberg DE, Lucas FL, Birkmeyer JD, Bredenberg CE, Fisher ES. Variation in carotid endarterectomy mortality in the Medicare population: trial hospitals, volume, and patients characteristics. *JAMA* 1998;279:1278-81.
17. Gupta PK, Pipinos II, Miller WJ, Gupta H, Shetty S, Johanning JM, et al. A population-based study of risk factors for stroke after carotid endarterectomy using the ACS NSQIP database. *J Surg Res* 2011;167:182-91.
18. American College of Surgeons, National Quality Improvement Program. Available at: <http://site.acsnsqip.org/>. Accessed June 21, 2012.
19. Vascular Study Group of New England. Available at: <http://www.vascularweb.org/regionalgroups/vsgnec/Pages/home.aspx/index.html>. Accessed June 22, 2012.
20. Pine M, Jordan HS, Elixhauser A, Fry DE, Hoaglin DC, Jones B, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;297:71-6.
21. Fry DE, Pine M, Jordan HS, Elixhauser A, Hoaglin DC, Jones B, et al. Combining administrative and clinical data to stratify surgical risk. *Ann Surg* 2007;246:875-8.

Submitted Jul 30, 2012; accepted Jan 1, 2013.

#### INVITED COMMENTARY

**Timothy F. Kresowik, MD, Iowa City, Iowa**

This report by Bensley et al emphasizes the important concept that databases are not all “created equal,” and thus, it is essential that the limitations of any database be recognized when drawing conclusions from analyses. Some of the differences identified in their report are a result of differences in definition between databases rather than “accuracy” per se. This has long been a problem for carotid procedures.

By far, the most important patient-related determinant of risk for a carotid intervention is symptom status. Problems arise from the historic division of symptom status into only two strata—symptomatic and asymptomatic—and that clinical reports and even randomized trials have used different definitions of symptomatic.

I submit that at least three levels of cerebrovascular symptom status and standardized definitions are necessary: asymptomatic

(never symptomatic in any distribution), symptomatic (recent, ipsilateral hemispheric or ocular symptoms), and other symptomatic (remote ipsilateral or any contralateral or global/vertebrobasilar symptoms). The use of these three strata and definitions reliably predict the risk of carotid endarterectomy.<sup>1</sup>

The current report clearly identifies some of the limitations of current administrative databases vis-à-vis clinical registries. I do not think, however, that one can assume that clinical registries, such as the National Surgical Quality Improvement Program or the Society for Vascular Surgery Vascular Quality Initiative, are feasible or even necessarily superior to administrative databases for all the potential uses of procedural outcome data. There is tremendous current pressure to have outcome performance measures available for accountability, such as public reporting or pay-for-performance